Documenting Intraspecfic Genome Size Variation in Soybean

A. Lane Rayburn,* D. P. Biradar, R. L. Nelson, R. McCloskey, and K. M. Yeater

ABSTRACT

Reports of genome size variation in soybean [*Glycine max* (L.) Merr.] have ranged from 40 to 0%. This wide range has resulted in doubts of the existence of intraspecific DNA variation in soybean. Eighteen soybean lines were analyzed by flow cytometry to determine their genome size. The lines were selected on the basis of diversity of origin. Intraspecific genome size variation was observed at approximately 4%. To ensure that the variation observed was reproducible and not due to technique error, the two highest and lowest genomes size accessions were reanalyzed. The order and variation observed between the high and low genome size accessions were maintained. To ensure further that the differences were reproducible, seeds from the two highest and lowest genome size accessions were planted in different locations in the USA, grown to maturity, harvested, and the seeds returned to Illinois. The harvested seed was analyzed and again the order and variation in genome size between the high and low genome size accessions were similar to the previous two analyses even though more than 1 yr had passed between the analysis. In addition, two experiments using *Amaranthus palmeri* S. Wats as an internal standard were conducted. In both of these experiments, the observed variation between the previously reported high and low genome size soybean lines was approximately 1 to 2%. The variation between the high and low genome size soybean lines is reproducible. The variation reported here indicates that the DNA amount variation is between 1 and 4%, lower than was originally reported.

Over the past several years, a controversy has been occurring within the literature with respect to intraspecific nuclear DNA content variation in plants. While certain species such as maize (*Zea mays* L.) have been excluded from the controversy, reported DNA content variation in many plants species has come under close scrutiny (Greilhuber, 1997, 1998). Serious doubts as to whether intraspecific DNA content variation exists and, if it exists, the extent of this variation. In no other species is this more apparent than in soybean.

Doerschug et al. (1978) first reported DNA content variation in soybeans. Upon examination of 11 soybean lines, the amount of DNA per nucleus was reported to range from 1.84 to 2.61 pg. This represents over a 40% variation in nuclear DNA content. Graham et al. (1994) observed a 15% variation among soybean cultivars while Rayburn et al. (1997) observed a 12% variation among 90 Chinese soybean introductions. Chung et al. (1998) observed among 12 soybean strains a 4.6% DNA content variation. However, Greilhuber and Obermayer (1997) and Obermayer and Greilhuber (1999) were not able to reproduce the reported DNA content variation in soybean and concluded that the soybean genome size was constant and that any reported DNA content variation was probably due to technical errors or experimental design flaws. Given the wide range of reported DNA variation in soybean, >40 to 0%, it is important to determine which level of variation is correct.

While various levels of variation are reported, the most incongruent variation is the 40% variation reported by Doerschug et al. (1978). This is also the only study using Feulgen cytophotometry exclusively. The remaining soybean studies referred to above all used flow cytometric techniques. According to Greilhuber (1998) suboptimal performance of the Feulgen reaction is one of the most common sources of reported intraspecific DNA variation. Given these observations, it is possible that the 40% DNA variation reported was not true DNA variation. In addition, if the DNA variation was technical in nature the lack of correlation between protein composition and nuclear DNA content could have been the result of technical noise masking the correlation. The objective of this study was to determine if intraspecific nuclear DNA variation truly exists in soybean.

MATERIALS AND METHODS

In the initial experiment, 18 lines were analyzed for nuclear DNA content (Table 1). Nine of the lines are U.S. cultivars, while the remaining lines are introductions from Japan, Russia, South Korea, and China. The three lines from Japan and two from China are released varieties while the others are more primitive landraces. For the remaining experiments, the two lines with highest DNA content (PI 227324 and PI 266085) and the two lines with the lowest nuclear DNA content (PI 253666 and PI 437088) were chosen.

The plants were grown and nuclei isolated and stained according to Rayburn et al. (1997). The fluorochrome propidium iodide was used to stain the nuclei. For Exp. 1, all 18 lines plus a Burlison standard were planted each day for 10 d. One week after planting, the hypocotyl portion of the stem of each line was removed and the nuclei isolated and stained. Burlison was used as an external standard. On each day of analysis, two samples of Burlison standard were run along with the 18 lines. The voltage of the photomultiplier tube (PMT) was adjusted each day so that the mean of the G1/G0 peak of the external standard was about 100 (Fig. 1). The data were then collected each day from the experimental lines without altering the PMT voltage, laser power, or flow rate. The nuclear DNA content of each line was determined relative to the mean DNA content of the two Burlison samples. The mean of the G1/G0 peak of the experimental line was divided by the mean of the two external standard lines then multiplied by 100.

For the second experiment, the four lines representing the

A. Lane Rayburn, R. McCloskey, and K. M. Yeater, Dep. of Crop Sciences, University of Illinois at Urbana-Champaign, 320 ERML, 1201 W. Gregory, Urbana, IL 61801, USA, D.P. Biradar, Department of Biotechnology, University of Agricultural Sciences, Dharwar 580 005 Karnataka, India, and R. L. Nelson, USDA-Agricultural Research Services, Soybean/maize Germplasm, Pathology, and Genetics Research Unit and Dep. of Crop Sciences, 1101 W. Peabody Drive, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
Received: 20 Feb. 2002.*Corresponding author (arayburn@uiuc.edu).

Published in Crop Sci. 44:261–264 (2004).

677 S. Segoe Rd., Madison, WI 53711 USA

Table 1. Genome size of the eighteen soybean lines analyzed.

Line	*N*	Genome size†
PI 266085 (China)	10	100.6 ± 1.8
PI 227324 (China culitvar)	10	100.3 ± 0.7
PI 507430 (Japan cultivar)	10	99.8 ± 1.6
Burlison	10	99.8 ± 0.9
PI 398516 (South Korea)	10	99.5 ± 1.3
Hodgson	10	99.1 ± 1.1
Hobbit	10	99.4 ± 2.0
Sloan	10	99.4 ± 1.9
Barc 8	10	99.3 ± 1.1
PI 92705 (China)	10	99.2 ± 1.5
Corsoy 79	10	98.5 ± 1.6
Barc 6	10	98.4 ± 1.6
PI 407788 (South Korea)	10	98.2 ± 1.9
Provar	10	97.9 ± 1.6
Pella	10	97.7 ± 3.1
PI 423948 (Japan cultivar)	10	97.7 ± 2.4
PI 253666 (China culitvar)	10	97.0 ± 1.1
PI 437088 (Russia)	10	96.9 ± 1.8
		LSD = 1.5**

** $\alpha = 0.01$.
† In Arbitrary Units (A.U.) with the standard defined as having 100 A.U.

nuclear DNA content extremes were planted on three separate days. After 1 wk, two samples were isolated per line per day. The nuclei were stained as described above. Burlison was again used as an external standard.

For the third experiment, the four lines described above were analyzed. However, the seeds used were grown at three different field locations. PI 253666 was grown in Illinois, Maryland, and North Carolina. PIs 227324, 437088, and 266085 were grown in Illinois, Minnesota, and Iowa. The seeds of all lines were harvested and shipped back to Illinois for nuclear DNA analysis. On each day, for 4 d, seeds were planted for each line at each location. Again, two samples of Burlison were used as external standards. One week after planting, one plant per line per location was analyzed.

In the above experiments, the nuclei were analyzed on a Coulter EPICS 750 series flow cytometer cell sorter (coulter Electronics, Hialeah, FL, USA). The excitation wavelength was 488 nm with 500 nuclei analyzed per sample. The G1/G0 peak was used for analysis (Fig. 1).

Two additional experiments were also run with the four lines described above. The seed used from the lines PI 227324, 437088, and 266058 was seed harvested in Iowa and Minnesota, while the seed from PI 253666 was harvested from Maryland and North Carolina. The seeds were planted and plants grown as previously described. In Exp. 4, 11 plants of each type from each state were analyzed. In Exp. 5, five plants from each accession from each state were analyzed for a total of 10 plants per PI accession. In these experiments, *A. palmeri* was used as an internal standard. Tissue from each soybean plant was harvested as previously described. Along with the individual hypocotyl portion of each plant a 13-mm fully expanded leaf of *A. palmeri* was placed in the same10 mL of extraction buffer. The tissues were chopped and ground together to release the nuclei. The nuclei were then stained together as described in Rayburn et al. (1997).

The nuclei were analyzed on a Coulter Epics XL (coulter Electronics). All the nuclei from each plant were coded and the samples run in a blind format. The data for each day were not decoded until ready for analysis. The means of the G1/G0 peaks were determined and the DNA amount of each soybean plant was reported relative to the internal standard {soybean DNA = [(G1/G0) mean of soybean/(G1/GO) mean *of A. palmeri*]}.

RESULTS AND DISCUSSION

The genome sizes of the 18 lines ranged from 100.6 to 96.9 Arbitrary Units (A.U.) (Table 1). Upon statistical analysis a significant difference was observed among the lines at $p < 0.0001$. LSD analysis indicated that only one of the named cultivars had a significantly different genome size. Thus there was little DNA variation within the cultivars.

Along with the seed lot of Burlison used as an external standard, a different seed lot of Burlison was used as an experimental cultivar. The mean of the standard was defined as 100 A.U. The raw data of each experimental line was divided by the mean of the two external standards and multiplied by 100. Therefore, the theoretical value of the experimental Burlison was 100 A.U (the same as the external standard). The actual value obtained in this study was 99.8 A.U.

The ≈4% variation among the soybean lines is due

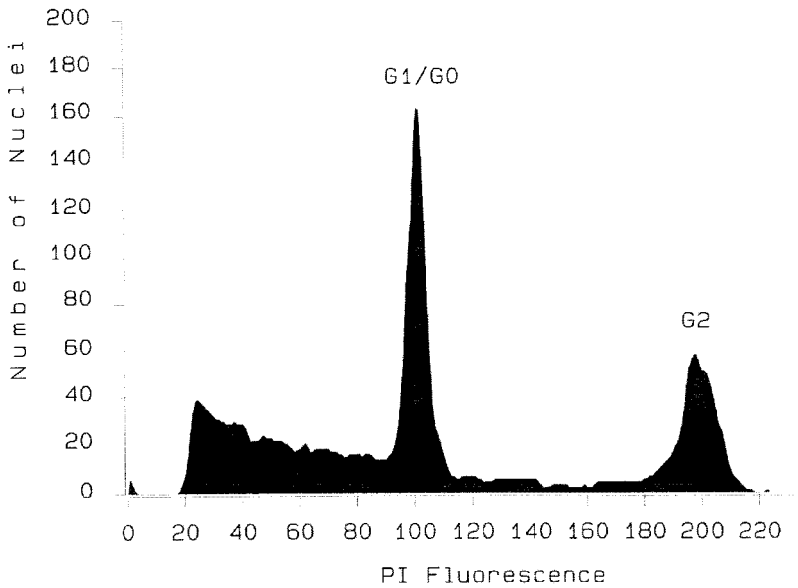

Fig. 1. Flow histogram of the soybean cultivar Burlison. The CV of the G1 peak is 2.96 and the G1/G2 ratio is 1.97.

Table 2. Pairwise comparison of highest and lowest genome size soybean in Exp. 2

Line	*N*	Mean	Standard deviation
PI 266085	6	100.1	1.0
PI 227324	6	99.3	1.5
PI 437088	6	97.6	1.6
PI 253666	6	94.8	3.5
Differences of least square means			
Line	Line	*t* Value	$P > 1$
PI 266085	PI 227324	−0.74	0.49
PI 266085	PI 437088	1.81	0.12
PI 266085	PI 253666	−3.36	0.02
PI 227324	PI 437088	1.07	0.33
PI 227324	PI 253666	2.63	0.04
PI 437088	PI 253666	−1.56	0.12

Table 3. Pairwise comparison of highest and lowest genome size soybean in Exp. 3.

Line	*N*	Mean	Standard deviation
PI 266085	12	99.0	2.5
PI 227324	12	99.1	2.2
PI 437088	12	97.7	2.5
PI 253666	12	92.7	3.0
Differences of least square means			
Line	Line	*t* Value	$P > 1$
PI 266085	PI 227324	0.09	0.93
PI 266085	PI 437088	0.79	0.45
PI 266085	PI 253666	−3.80	0.005
PI 227324	PI 437088	0.87	0.41
PI 227324	PI 253666	3.89	0.004
PI 437088	PI 253666	−3.01	0.02

to the variation among the non-U.S. plant introductions. The low genome size variation observed in this study more closely agreed with the other flow cytometric studies than with the microspectrophotometry study. Indications are that the 40% variation observed in Doerschug et al. (1978) could be due to technical error. It should be pointed out, however, that since different cultivars were used in each study and the studies were conducted over 23 yr apart, the differences could indeed have been real.

In the second experiment, four lines were selected for analysis. PI 266085 and PI 227324 had the highest genome sizes in the previous experiment, while PI 437088 and PI 253666 had the smallest genome sizes. Seeds from the same seed lots were used. In this experiment, PI 266085 and PI 227324 again had the highest genome sizes while PI 437088 and PI 253666 had the smallest genome sizes (Table 2). Upon statistical analysis, PI 253666 had a significantly lower genome size than either PI 266085 or PI 227324 (Table 2). Although the genome size of PI 437088 was not significantly different from either PI 266085 or PI 227324, no significance was observed between the genome sizes of PI 437088 and PI 253666.

The results of Exp. 1 are in agreement with Exp. 2; significant genome size variation was observed among the soybean lines. The same two lines had the largest and smallest genome sizes respectively. The range between the two highest and lowest genome size lines remained ≈4% in both experiments. Although the same trends were noted, slight discrepancies were still observed. It was therefore decided to run a third experiment.

Seeds for Exp. 3 were obtained by growing the original seed lots out in three distinct locations, harvesting seeds from those plants, and analyzing the seedlings. Thus, any potential biological variation would be maximized. In addition, this experiment was performed well over 1 yr from the first experiment. This time frame could allow for increased technical and instrumentation variability. As in the previous two experiments, PI 266085 and PI 227324 again had the highest genome sizes while PI 437088 and PI 253666 had the smallest genome sizes (Table 3). Upon statistical analysis, PI 253666 had a significantly lower genome size than either PI 266085 or PI 227324 (Table 3). Also the genome size of PI 437088 was not significantly different from either PI 266085 or PI 227324; however, unlike the previous two experiments, significance was observed between the genome sizes of PI 437088 and PI 253666.

Regardless of any potential biological, technical, or instrument induced genome size variation, the ranking of the lines in Exp. 3 was identical to the rankings previously observed. The percentage difference between the largest and smallest genome sized lines was ≈4.0%. Significant genome size variation among the soybean lines also substantiated the earlier results.

Recently concerns have been raised about the presence of PI fluorescence inhibitors (FIs) and their effect on flow cytometric determination of intraspecific DNA content variation (Price et al., 2000). Plants that have the same genome size but differ in the presence of FIs may appear to have different genome sizes. To address these concerns Exp. 4 was conducted. By co-chopping a leaf of *A. palmeri* with the soybean the *A. palemeria* acts as an internal standard. Price et al. (2000) reported that cochopping with an internal standard appears to compensate for the presence of FIs. Thus if any FIs are present in soybean, the use of an internal standard should compensate for them. The G1/G0 peaks of *A.*

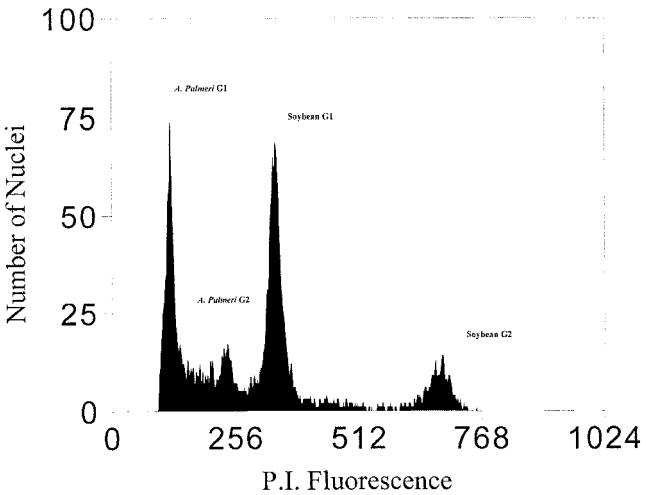

Fig. 2. Flow histogram of Soybean PI accession 266085 co-chopped and analyzed with *A. plameri* as an internal standard. The mean of the G1/G0 peak of the Soybean line was 338.6 and the mean of the internal standard was 117.4. The relative DNA amount of the experimental standard sample was 2.884 times the internal standard.

Table 4. Genome size of the soybean lines, relative to the internal standard, analyzed in Exp. 4.

Line	*N*	Relative mean†	Standard deviation
PI 266085	22	2.847	0.093
PI 227324	22	2.825	0.064
PI 437088	22	2.814	0.071
PI 253666	22	2.814	0.076
	LSD = 0.031*		

* Alpha = 0.05.
† Mean of the G1/G0 soybean peak divided by the mean of the G1/G0 internal standard peak.

palmeri and soybean were distinct and as such the mean of each of these peaks were easily discernable (Fig. 2). PIs 227324 and 266085 had the highest DNA amounts with PIs 437088 and 253666 having the lowest DNA amounts (Table 4). This is the identical trend observed in the external standard studies.

Upon statistical analysis, both PI 253666 and PI 437088 had a significantly lower DNA amount than PI 266085 (Table 4). A DNA amount difference of ≈1.0% was observed between the high and low DNA lines in this experiment. Although the two high DNA content lines and the two low DNA content lines were consistent, the amount of variation between the high and low DNA lines was less than the previous experiments.

To confirm the internal standard results, Exp. 5 was conducted also with the internal standard. As in all of the previous four experiments, the two lines with the highest reported genome sizes did indeed have the highest genome size and the lines reported to have the lowest genome size both were observed to have the lowest genome size with ≈2% variation observed between the averages (Table 5). Both PI 437088 and PI 253666 had statistically significantly lower DNA amounts than did PIs 266085 and 227324.

In conclusion, all five experiments agreed with respect to the ranking of the two highest and two lowest genome size lines. At no time did one of the two lines with the lowest genome size ever rank above the two lines with the largest genome size. In addition, a significant genome size variation was observed in all five experiments. Therefore, little doubt remains as to the existence of intraspecific genome size variation in soybean. The results of this study indicate that the 40% variation may be an overestimate of the DNA variation. The 1 to 4% variation observed in this study is consistent with Chung et al. (1998). Such small variation among lines within plant species has been previously reported and cytologically confirmed (Wetzel et al., 1999; Wetzel and Rayburn, 2000). Since soybean is a self-pollinating species, one might expect it to have a lower amount of intraspecific DNA variation than an outcrossing plant species (Bennett et al., 2000). As stated by Rayburn et al. (1997), genome size variation exists in soybeans and appears much less than the genome size variation observed in maize.

Table 5. Genome size of the soybean lines, relative to the internal standard, analyzed in Exp. 5.

Line	*N*	Relative mean†	Standard deviation
PI 266085	10	2.892	0.058
PI 227324	10	2.883	0.046
PI 437088	10	2.820	0.073
PI 253666	10	2.826	0.075
	LSD = 0.058*		

* Alpha = 0.05.
† Mean of the G1/G0 soybean peak divided by the mean of the G1/G0 internal standard peak.

ACKNOWLEDGMENTS

The authors thank Dr. B. Pilas of the Flow Cytometry Facility, a resource of the University of Illinois Biotechnology Center, for her assistance. Funding for this research was provided in part by the Illinois Soybean Program Operating Board grant No. 94-20-146-3.

REFERENCES

Bennett, M.D., S. Johnston, G.L. Hodnett, and H.J. Price. 2000. Allium cepa L. cultivars from four continents compared by flow cytometry show nuclear DNA constancy. Ann. Bot. 85:351–357.

Chung, J., J.H. Lee, K. Arumuganathan, G.L. Graef, and J.E. Specht. 1998. Relationship between nuclear DNA content and seed and leaf size inn soybean. Theor. Appl. Genet. 96:1064–1068.

Doerschug, E.B., J.P. Miksche, and R.G. Palmer. 1978. DNA content, ribosomal-RNA gene number, and protein content in soybeans. Can. J. Genet. Cytol 20:5331–5338.

Graham, M.J., C.D. Nickell, and A.L. Rayburn. 1994. Relationship between genome size and maturity group in soybean. Theor. Appl. Genet. 88:429–432.

Greilhuber, J. 1997. The problem of variable genome size in plants (with special reference to woody plants). p. 13–34. *In* Z. Borzan and S.W. Schlarbaum (ed.) Cytogenetic studies of forest trees and shrub species. Croatian Forests, Inc., Faculty of Forestry, University of Zagreb, Croatia.

Greilhuber J. 1998. Intraspecific variation in genome size: A critical reassessment. Ann. Bot. 82(Supplement A): 27–35.

Greilhuber, J., and R. Obermayer. 1997. Genome size and maturity group in Glycine Max (soybean). Heredity 78:547–551.

Obermayer, R., and J. Greilhuber. 1999. Genome size in Chinese soybean accessions– Stable or variable? Ann. Bot. 84:259–262.

Price, H.J., G. Hodnett, and J.S. Johnston. 2000. Sunflower (Helianthus annus) Leaves contain compounds that reduce nuclear propidium iodide fluorescence. Ann. Bot. 86:929–934.

Rayburn, A.L., D.P. Biradar, D.G. Bullock, R.L. Nelson, C. Gourmet, and J.B. Wetzel. 1997. Nuclear DNA content diversity in Chinese soybean introductions. Ann. Bot. 80:321–325.

Wetzel, J.B., S. Aref S, V.C. Baligar, A.L. Rayburn. 1999. A lack of nuclear DNA content variability among wheat isolines differing in aluminum response. Ann. Bot. 83: 725–728.

Wetzel, J.B., and A.L. Rayburn. 2000. Use of fluorescence genomic in situ hybridization (GISH) to detect the presence of alien chromatin in wheat lines differing in nuclear DNA content. Cytometry 41:36–40.